SINGLE EXPOSURE VS TONE MAPPED HIGH DYNAMIC RANGE IMAGES: A STUDY BASED ON QUALITY OF EXPERIENCE

Manish Narwaria, Matthieu Perreira da Silva, Patrick Le Callet, Romuald Pepion

SINGLE EXPOSURE VS TONE MAPPED HIGH DYNAMIC RANGE IMAGES: A STUDY BASED ON QUALITY OF EXPERIENCE

Manish Narwaria, Matthieu Perreira Da Silva, Patrick Le Callet, Romuald Pepion

LUNAM University - IRCCyN CNRS UMR 6597, 44306, Nantes, France

ABSTRACT

Tone mapping operators (TMOs), employed to fit the dynamic range of High Dynamic Range (HDR) visual signals to that of the display, are generally non-transparent and modify the visual appearance of the scene. Despite this, tone mapped content generally tends to have more visual details as compared to a single exposure scene. It is however not clear if the extra details in tone mapped HDR affect user preferences over a single exposure content in terms of scene appearance and to what extent. This paper aims to shed light on this issue via a comprehensive subjective study. Our results reveal that there is no statistical evidence to establish if the users preferred tone mapped content over the single exposure version as closer representation of the corresponding HDR scene. We present those results as well as outline the possible factors contributing to this somewhat unexpected finding.

Index Terms— Quality of Experience (QoE), High Dynamic Range (HDR), tone mapping

1. INTRODUCTION

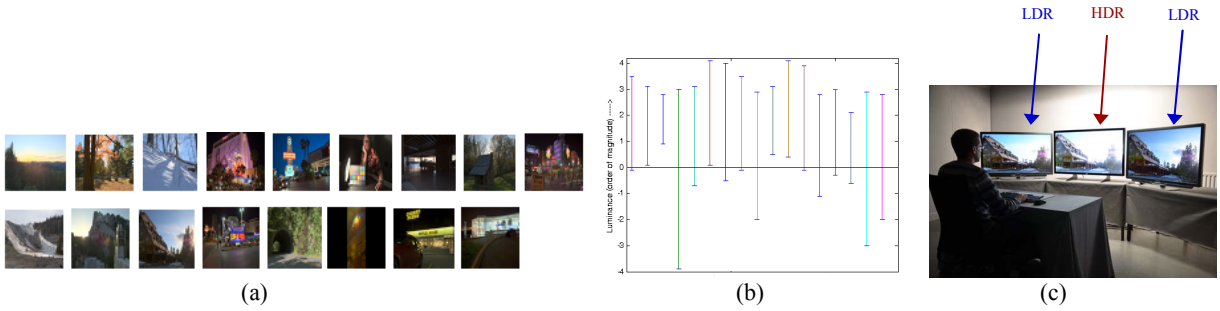
High Dynamic Range (HDR) has emerged as an exciting research area in the quest for providing truly immersive quality of experience (QoE) to the users [1]. This is achieved by expanding the capture and display capabilities to match the instantaneous human vision range which in turn results in high contrast, compelling visual signals. HDR has also attracted a lot of recent research attention from the signal processing community for tasks such as HDR compression, processing, enhancement and quality assessment.

While algorithmic approaches to HDR creation (such as multi-exposure fusion) provide a practically feasible solution for HDR content generation [1], displaying HDR on LDR displays remains a challenge. Since the cost of HDR display technologies is currently quite high and yet to reach consumer levels, the only alternative is to display HDR content directly on commonly available LDR devices such as CRT, LCD displays, printers etc. To this end, tone mapping operators (TMOs) are often employed to tackle the mismatch between the dynamic range of HDR content and LDR displays. The underlying principle of TMOs is to retain perceptually significant details in order to achieve range reduction. As a result of this, tone mapped LDR content tends to have more details in comparison to a single exposure photograph. However, to our knowledge, there are no formal studies to establish if these extra visual details improve the visual appearance and eventually make the tone mapped content a realistically reasonable substitute for HDR. To answer this and obtain further well-grounded insights, we present the results from a subjective study that we carried out. Based on the statistical analysis, the main conclusion from our study is that observers did not consistently prefer tone mapped over single exposure. We further explain and analyze our findings based on the thesis that visual appearance of images is dependent on several factors including details, color appearance, naturalness etc.

2. MOTIVATION AND RELATION TO PRIOR WORK

Several studies in the past have examined the impact of TMOs on the visual quality of HDR content [2]. The general philosophy of a majority of them is to display tone mapped HDR content on LDR displays and ask observers to rate and/or compare them. Apart from obtaining user preference, many of these studies also instructed the observers to provide further information on perceptual attributes such as image contrast, apparent level of detail, and apparent naturalness etc. The reader is referred to [2] [3] for quick summary of these studies. There are however two major drawbacks associated with the existing studies.

First, none of them compared tone mapped and single exposure content. Such comparison, in our opinion is meaningful because it can provide valuable insights into whether the TMOs bring any added value or not pertaining to the visual experience of the observers. While TMOs, in general, can retain more details they can also affect perceptual attributes such as color/luminance, naturalness overall contrast etc. As a result, it is not proper to assume that tone mapped content will always be preferred over a single exposure one. Second, it is worth considering is that with the exception of few studies (eg. [4]), none display the HDR content to the observers. Thus, observers could only select a better looking tone mapped image (or video) based



(a) (b) (c)

Fig. 1. (a) Reference test content (these are tone mapped versions) used in the study. First row: C1 to C9, Second row: C10 to C17, (b) Luminance range of content in order of magnitudes (left to right C1 to C17), (c) Experimental setup with the HDR display in the center and the two same type LDR displays on left and right. Figure best viewed in color.

on 'abstract realism'. The absence of an HDR reference may lead to less accurate characterization of TMOs with respect to how they modify HDR and what visual features are damaged. Indeed, the study in [13] found that with access to a reference scene, the observers' opinion can change. To alleviate the mentioned drawbacks associated with existing studies, we included single exposure (in addition to tone mapped content) images for the observers to vote. Further, we employed an HDR display to provide the observers a reference to compare with. As a result, we could examine which LDR condition was able to represent HDR better.

3. DETAILS OF THE STUDY

In this section we provide a brief description of the test scenes and experimental setup.

3.1. Test content and TMOs

Our study involved 17 HDR content (denoted as C1 to C17) which were obtained from [5] and their tone mapped versions are shown in Figure 1 (a). These include a wide range of content and their luminance ranges in terms of orders of magnitude is shown in Figure 1 (b). To generate the single exposure content, we adjusted the exposure values to approximately mid level for each content separately. For tone mapping, we chose 3 TMOs namely linear TMO (based on saturating 2.5% of the total number of pixels for both low and high luminance and then applying gamma correction with gamma value being 2.2)), iCAM06 [6] and the global TMO proposed by Reinhard et al. [7]. Thus for each HDR content we obtained four corresponding LDR content (i.e. 4 LDR conditions).

3.2. Experimental setup

Observers were seated in a standardized room conforming to the International Telecommunication Union Recommendation (ITU-R) BT500-11 recommendations [8]. For displaying the HDR content, SIM2 HDR47E S 4K display [9] was used, which is a 47-inch, 1080p LCD TV with maximum displayable luminance of 4000 cd/m². The two LDR displays were 46-inch (Philips 46PFL9705H) with maximum displayable luminance of 200 cd/m². The viewing distance was set to approximately three times the height of the screen (active part). For each comparison, the observers saw three stimuli: one on the HDR display placed at the center and two on the LDR displays on either side as shown in Fig. 1 (c). Since there are two types of displays, the room illumination was adjusted accordingly. In particular with HDR display (brighter) in the center, the illumination at the center (just above the HDR display) was set to 100 cd/m² while the diffused light (about 50 cd/m²) made up the illumination for each LDR display. Such a setup ensured a suitable illumination setting for the observers, and they were comfortable while viewing both HDR and LDR stimuli. A paired comparison (PC) methodology was adopted and the observers were instructed as "Please choose the image (left or right) that is more similar to the reference image (center)". To avoid bias effects, the order of displaying LDR stimuli was randomized i.e. the single exposure and tone mapped content appeared on either of the LDR displays randomly. A total of 38 observers participated in the study and they had normal or corrected to normal visual acuity and normal color perception. The observers were also asked for the reason for discarding the non-selected image. The question was "Why did you discard this image?". To answer this, they had 3 choices: low fidelity of colors / luminance, loss of details, lack of naturalness. Since all the observers were naive for the purpose of this (not expert in image or video processing) study, the physical meaning of each of these choices was described in details on a separate sheet during the experiment. Further, only one choice was allowed to ensure that the most significant factor affecting observers' decision was obtained.

4. QUANTITATIVE ANALYSIS OF EXPERIMENTAL RESULTS

The outcome of the PC data can be best visualized in the form of a preference matrix whose dimension, in our case, will be four by four (corresponding to four LDR stimuli per reference HDR content). In this paper, we will use the notations I, L, R and S respectively denote iCAM06, linear,

	S	R	L	I	Total
S	–	345	310	331	986
R	301	–	323	312	936
L	336	323	–	276	935
I	315	334	370	–	1019

Table 1. Overall Preference matrix

Reinhard and single exposure. The overall preference matrix across content is given in Table 1 in which each value denotes the number of times the LDR stimuli for the LDR condition mentioned in row was perceived as being closer, in overall similarity, to the reference HDR in comparison to the one in column. For instance, the fourth cell value of 331 indicates that the single exposure was preferred 331 times over iCAM06 across content. The last column in Table 1 shows the total number of votes in favor of the corresponding LDR condition and is obtained by summing the entries row wise. The reader will notice that the diagonal entries are not considered and this is because an LDR stimulus was never compared to itself. Also it can be seen that the aggregate votes (across content and LDR conditions) is 3876 which is equal to the product of number of comparisons for each content ($^4C_2$), total number of observers and the total number of content i.e. $6 \times 38 \times 17$.

4.1. Overall Analysis

A quick look at the total number of votes (refer to Table 1) for each LDR condition indicates that the observer preference is fairly scattered i.e. there is less agreement between observers on which LDR stimuli best represented the reference HDR stimuli. A further glance at the last column of Table 1 reveals that all the 4 types of LDR content got a similar number of votes in the overall scenario. This clearly indicates that the tone mapped content was not necessarily preferred over the single exposure one despite having more visual details. The subjective rankings were further analyzed based on the Bradley Terry (BT) model [11] which is a statistical model that can provide relative distances between stimuli by scaling the preferences. We have presented in Figure 2 (a), the results from BT analysis for the overall scenario (across content). The x-axis shows the LDR condition (I, L, R and S) while the y-axis represents the BT score (higher implies better) and error bars denote the 95% confidence intervals. One can observe that the single exposure (along with iCAM06) is preferred (both have higher BT scores) over Reinhard and linear TMO and the single exposure LDR condition is the second best in terms of closer appearance to HDR. Further statistical results from the BT model are presented in Fig. 2 (b) in which black color cell means that the corresponding LDR condition is statistically indistinguishable and white denotes the opposite case. We can see that iCAM06 is statistically better then Reinhard and linear TMO but indistinguishable from the single exposure. In fact, the first row of Figure 2 (b) (all cells being black) provides statistical evidence that, overall, the observers perceived the single exposure LDR image as similar to the tone mapped LDR in terms of fidelity with HDR. This is a rather surprising result at first since we expected that tone mapped content should have been judged as closer to HDR given the emphasis TMOs place on preserving details. In the next section, we quantify the levels of details from a quantitative angle to facilitate further analysis and discussion.

4.2. Quantifying visual details with sharpness

Visual inspection of the LDR content clearly revealed that, in general, the tone mapped content preserved more visual details and overall appeared sharper as compared to the single exposure content. To further quantify the amount of visual details, we measured the sharpness of the LDR content objectively using two methods: (a) a simple gradient based method and (b) the method proposed in [10] (referred to as S3). The resulting sharpness values from the two methods have been shown in Figure 4 (a) and (b) respectively (a larger bar size for an LDR condition implies higher sharpness). These values further confirm that the tone mapped content was generally sharper than the single exposure. One can notice that the single exposure (denoted by S), generally, has the lowest sharpness (as indicated by the smallest bar Figure 4). Therefore, despite the tone mapped content being sharper in general, the observers did not always select it. We explain this by considering other factors that possibly contribute to scene appearance in the next section.

5. QUALITATIVE ANALYSIS OF PERCEPTUAL ATTRIBUTES

The quantitative analysis presented in the previous section confirms that difference in preference between single exposure and tone mapped content is statistically insignificant. In fact, in many cases for each content separately, the former was preferred over the latter (to re-iterate, the results reported in Figure 2 (a) and (b) are across content). As already mentioned, this is surprising especially due to the fact that tone mapped content in general had more details (this has been objectively confirmed in section 4.2). In the following, we present the analysis to explain these observations.

5.1. Context of visual details

Even though the global sharpness tends to be higher for tone mapped content, it is useful to analyze it from local perspective. The reason is that the context (or location) of details is also important. That is, it might be possible that a TMO preserves details in areas that may not be noticed

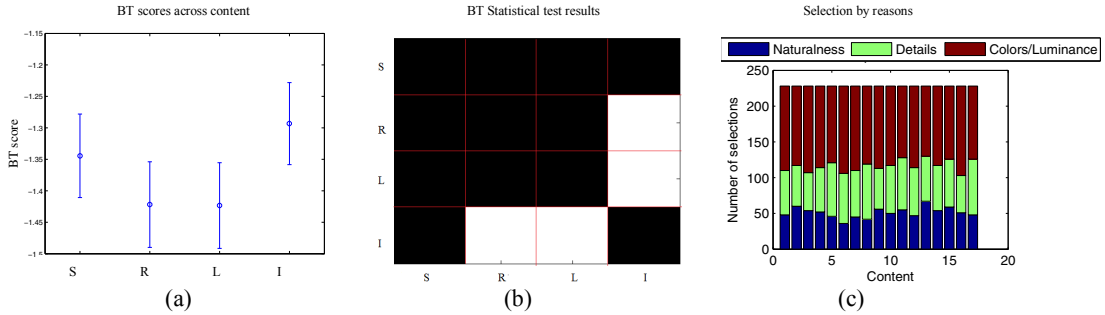


Fig. 2. Quantitative results, (a) Average preference scores (based on BT analysis) across all the content with the error bars denoting the 95% confidence intervals, (b) Results from BT statistical test (black cell means that there is no statistical difference between the row and column LDR and white cell represents the opposite case), (c) Number of selections based on each of three reasons (naturalness, details and color/luminance) for each content. Figure best viewed in color.

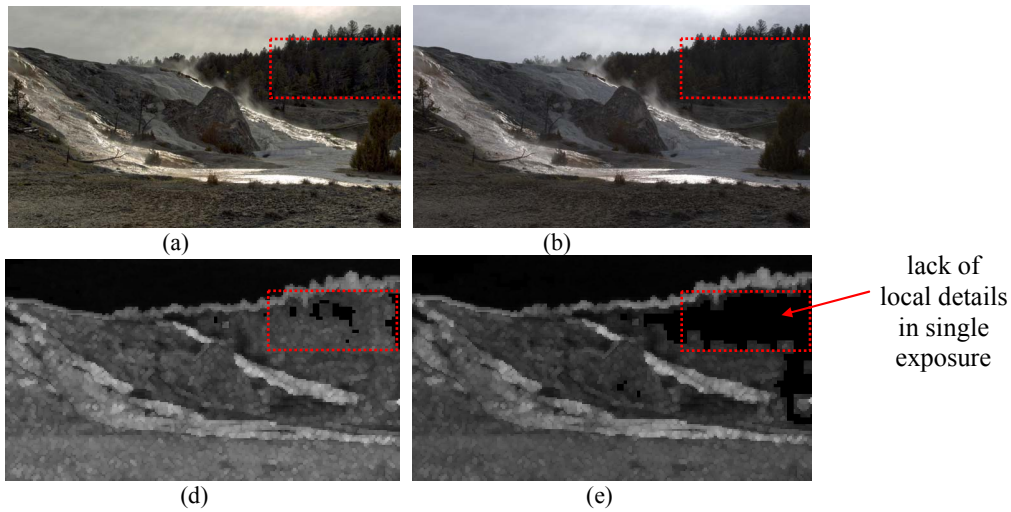


Fig. 3. Local sharpness, (a) iCAM06 LDR, Sharpness score = 0.0895, (b) Single exposure LDR, Sharpness score = 0.0233, (d) and (e) are respectively the local sharpness maps of (a) and (b). Figure best viewed in color.

by observers (for instance in the background) or the visibility of the details is masked. To illustrate this point further, an example is given in Figure 3 where we have shown the local sharpness maps (based S3 method [10]) in the second row for the single exposure and iCAM06 LDR stimuli (the corresponding reference HDR stimuli is content C10 for which there was no statistically preferred LDR condition). Note that in these local sharpness maps, brighter color implies more sharpness. Further, the overall (i.e. global) sharpness scores have also been indicated in the figure caption (higher score indicates more sharpness). In these images, we have also highlighted (by a red box) a local area where more details have been preserved by iCAM06 TMO in comparison to the single exposure content. One can notice that the highlighted area belongs mainly to the background. As a result, the presence of more details was less or not noticed by an average observer. Therefore, the local context of where details appear is also important. Apart from that, the characteristic of the details themselves can affect the user preference (for example observers may not prefer unrealistic or over enhanced details). Thus, the single exposure content was not necessarily discarded in spite of having lesser visual details.

5.2. Role of other perceptual attributes

Recall that the observers were not only asked to select the preferred LDR stimuli but also had to indicate one out of the three reasons (color/luminance, details and naturalness) for discarding the other. The supplementary question regarding the reasons for rejecting a particular LDR content provides a tractable way to analyze the findings of this study. Recall (refer to Table 2 and Figure 2 (a) and (b)) that the tone mapped LDR content was not statistically closer to HDR than the single exposure content. We know that TMOs generally retain more details by an appropriate (perceptually related in many cases) local (or global) transformation. However, this can tamper with other content (or scene) characteristics and in particular, naturalness and overall color/luminance could be modified. The observer preference

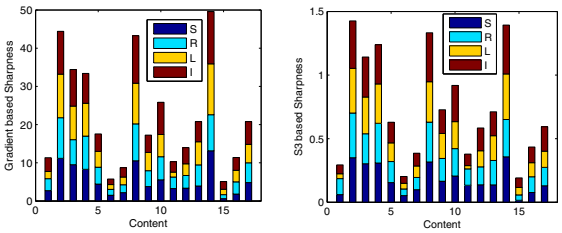


Fig. 4. Objective sharpness plots. (a) Gradient based sharpness, (b) S3 based sharpness. In both cases, higher value implies more sharpness. The legend refers to the 4 LDR conditions namely I, R, L and S. Figure best viewed in color.

information based on the three mentioned reasons is presented in Figure 2 (c) per reference HDR content. We find that color/luminance information had a larger effect (in several cases nearly 50% selections were made based on this attribute) on the observers' decision to reject a particular LDR content as compared to naturalness or the level of details. This suggests that the observers paid particular attention to the global color/luminance match while deciding which LDR content was closer to the reference HDR. Nevertheless, the other two factors namely details and naturalness also contributed in user preference. Thus, the decision on which LDR content is closer to HDR is complex and can involve a host of factors including the ones mentioned.

The results from the supplementary information shown in Figure 2 (c) also reveal that details (which many TMOs focus on) was not the most crucial factor. This explains, at least partly, why tone mapped content was not always preferred to single exposure. These findings also indicate that the TMOs in general should at least take into account factors like color/luminance and naturalness preservation and not just emphasize on preserving details. Moreover, even the detail preserving mechanisms in TMOs should be able to take into account the local context. Out of the 4 LDR conditions, iCAM06 TMO actually takes into account color information (and also local details) for tone mapping. This, in turn, explains the relatively better overall performance of this TMO (but statistically indistinguishable from single exposure) in producing images that were closer to the reference HDR.

6. CONCLUDING REMARKS

We studied if the extra visual details in tone mapped HDR content a played role in users' decision of selecting it as being closer to the HDR scene, in comparison to the single exposure image. The results indicate that this is not the case. We further provided insights and analysis and found that color/luminance, naturalness also play a role in user preference. Statistical analysis further confirmed that overall there was no significant differences between tone mapped and single exposure content. Our future study will focus on further analysis of the subjective data especially from the view point of visual attention which has been shown to be significantly modified by TMOs [12]. We will also investigate into how the characteristic of content affects its response to TMOs.

7. REFERENCES

[1] F. Banterle , A. Artusi, K. Debattista and A. Chalmers, "Advanced High Dynamic Range Imaging: Theory and Practice", ISBN: 978-156881-719-4, AK Peters (*CRC Press*), Natrick, MA, USA.

[2] M. Narwaria, M. Silva, P. Callet and R. Pepion "Effect of Tone Mapping on Visual Attention Deployment", *SPIE Conference on Applications of Digital Image Processing XXVII*, *vol. 8499,* 2012.

[3] M. Cadík, M. Wimmer, L. Neumann and A. Artusi, "Evaluation of HDR tone mapping methods using essential perceptual attributes", in *Computers & Graphics*, pp.330-349, 2008.

[4] P. Ledda , A. Chalmers, T. Troscianko, H. Seetzen, "Evaluation of tone mapping operators using a high dynamic range display", in *ACM Transactions on Graphics*, vol. 24, no. 3, pp. 640-648, 2005.

[5] http://www.cis.rit.edu/fairchild/HDRPS/HDRthumbs.html

[6] J. Kuang, G. Johnson and M. Fairchild, "iCAM06: A refined image appearance model for HDR image rendering", *Journal of Visual Communication and Image Representation*, vol. 18, no. 5, pp. 406-414, 2007.

[7] E. Reinhard, M. Stark, P. Shirley and J. Ferwerda, "Photographic tone reproduction for digital images" in *ACM Transactions on Graphics*, vol. 21, no. 3, pp. 267-276, 2002.

[8] Recommendation ITU-R BT.500-13, "Methodology for the subjective assessment of the quality of television pictures", 2012.

[9] http://www.sim2.com/HDR/

[10] C. Vu, T. Phan, and D. M. Chandler, "S3: A Spectral and Spatial Measure of Local Perceived Sharpness in Natural Images," *IEEE Transaction on Image Processing*, vol. 21, no. 3, 2011.

[11] R. Bradley and M. Terry, "Rank analysis of incomplete block designs: I. The method of paired comparisons, *Biometrika*, vol. 39, pp. 324 - 345.

[12] M. Narwaria, M. Silva, P. Callet and R. Pepion, "Tone mapping Based High Dynamic Range Compression: Does it Affect Visual Experience?", *Signal Processing: Image Communication*, vol. 29, no. 2, 2014.

[13] M. Ashikhmin and J. Goyal, "A Reality Check for Tone-mapping Operators", *ACM Transactions on Applied Perception*, vol. 3, no. 4, 2006.